by reference). Other examples of statistical analysis software available for principal-component-based methods include SPSS (SPSS Inc., Chicago, IL), JMP (SAS Inc., Cary NC), Stata (Stata Inc., College Station, TX) and Cluster.

In the Claims

Please delete claims 27-54 and 59-64 without prejudice. Applicants reserve the right to file divisional applications directed to these claims withdrawn from consideration in the present application.

Please amend the claims as follows.

1. (Amended) A method of predicting a biological activity of a test compound, comprising:

obtaining spectral data for the test compound and for a training set of compounds having known biological activities;

scaling the spectral data of the training set of compounds prior to deriving a pattern of spectral data associated with the biological activity;

deriving the pattern of spectral data associated with the biological activity from the spectral data of the training set of compounds; and

predicting the biological activity of the test compound by detecting similarities between the pattern of spectral data associated with the biological activity and a pattern of spectral data for the test compound.

2. (Amended) The method of claim 1, wherein the spectral data are obtained without first correlating the spectral data with corresponding structural features.

3. (Amended) The method of claim 1, wherein the pattern of spectral data associated with a biological activity is derived without first correlating the spectral data with corresponding structural features.

4. (Amended) The method of claim 1, wherein the pattern of spectral data of the training set is a pattern obtained by segmenting the spectral data of the training set of compounds into sub-spectral units.

5. (Amended) The method of claim 4, wherein the pattern of spectral data of the test compound is obtained by segmenting the spectral data of the test compound into substantially the same sub-spectral units into which the spectral data of the training set is segmented.

6. (Reiterated) The method of claim 1, wherein the spectral data is one type of spectral data.

C4 SUB. DI

7. (Amended) The method of claim 6, wherein the spectral data is one of nuclear magnetic resonance, mass spectral, infrared, ultraviolet-visible, fluorescence, or phosphorescence data.

8. (Reiterated) The method of claim 1, wherein the spectral data is a composite of different types of spectral data.

C10 SUB. DI

9. (Amended) The method of claim 8, wherein the composite comprises two or more of the group consisting of nuclear magnetic spectroscopy (NMR), mass spectroscopy (MS), infrared (IR) spectroscopy, and ultraviolet-visible (UV-Vis) spectroscopy.

10. (Amended) The method of claim 1, wherein the spectral data of the test compound is segmented into substantially the same spectral sub-units as the spectral data of the training set of compounds to produce the spectral pattern for the test compound.

11. (Amended) The method of claim 1, wherein scaling comprises auto-scaling.

12. (Amended) The method of claim 1, further comprising weighting the spectral data of the training set to emphasize signals that are important for determining an endpoint class of compounds in the training set before deriving the pattern associated with the biological activity.

13. (Amended) The method of claim 12, wherein weighting comprises Fisher-weighting.

14. (Amended) The method of claim 1, wherein detecting similarities between the pattern of spectral data associated with a biological activity of the training set and the pattern of spectral data for the test compound comprises statistical pattern recognition.

15. (Amended) The method of claim 10, wherein detecting similarities between the pattern of spectral data associated with a biological activity of the training set and the pattern of spectral data for the test compound comprises detecting relative intensities of one or more of the sub-units of the pattern of spectral data derived from the training set, and detecting relative intensities of signals associated with the same one or more sub-units of a spectrum of the test compound.

16. (Reiterated) The method of claim 15, wherein the relative intensities are canonical variate factors of the spectral data associated with a biological activity of the training set and the spectral signals of the test compound.

17. (Amended) The method of claim 1, wherein the method comprises artificial intelligence pattern recognition.

18. (Amended) A computer implemented method for predicting a biological activity of a test compound, comprising:

receiving spectral data for a test compound as input;

receiving spectral data and endpoint data of a training set of compounds having known biological activities as input;

segmenting the spectral data of the training set of compounds into sub-spectral units;

scaling the segmented spectral data of the training set of compounds;

detecting a pattern of spectral data associated with the biological activity; and

predicting the biological activity of the test compound by comparing the pattern of spectral data associated with the biological activity to the spectral data of the test compound to determine whether the spectral data of the test compound is similar to the spectral pattern associated with the biological activity of the training set and the test compound is predicted to share the biological activity.

19. (Amended) The computer implemented method of claim 18, wherein comparing comprises comparing with a statistical pattern recognition program.

20. (Amended) The computer implemented method of claim 19, wherein the spectral data for the test compound is segmented into substantially identical sub-spectral units as the training set spectral data, so that a signal within an individual sub-spectral unit is compared to the corresponding sub-spectral unit of the pattern.

21. (Amended) The computer implemented method of claim 18, wherein the spectral data are selected from the group consisting of nuclear magnetic resonance data, mass spectral data, infrared data, ultraviolet-visible data, fluorescence data, phosphorescence data, and composites of two or more such spectral data.

22. (Amended) The computer implemented method of claim 21, wherein the pattern associated with the biological activity of the training set is a set of canonical variate factors, and the spectral data for the test compound are compared to the canonical variate factors of the training set spectral data.

23. (Amended) The computer implemented method of claim 22, wherein the biological activity is binding affinity to a hormone receptor, and the canonical variate factors include peaks in sub-spectral units that are associated with hormone receptor binding of a pre-selected affinity.

24. (Amended) The computer implemented method of claim 23, wherein the spectral data comprise nuclear magnetic resonance data and mass spectral data.

25. (Reiterated) A computer readable medium having stored thereon instructions for performing the actions of claim 1.

26. (Reiterated) A computer readable medium having stored thereon instructions for performing the actions of claim 18.

SUB C12 D1

55. (Amended) The method of claim 1, wherein the spectral data comprises calculated spectral data.

56. (Amended) The computer implemented method of claim 18, wherein the spectral data comprises calculated spectral data.

57. (Amended) The computer implemented method of claim 56, wherein the calculated spectral data comprises calculated nuclear magnetic resonance data.

58. (Amended) The computer implemented method of claim 57, wherein the calculated nuclear magnetic resonance data comprises calculated ^{13}C NMR data.

SUB D1 C13

65. (Amended) The method of claim 1, further comprising:

predicting a second biological activity of the test compound by comparing the spectral data of the test compound to a second pattern of spectral data associated with the second biological activity to determine if the test compound shares the second biological activity, the second pattern derived using scaled spectral data and known endpoints of a training set of compounds for the second biological activity.

Please add the following new claims 66-90.

--66. (New) The method of claim 1, wherein the spectral data comprises ^{1}H, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{31}P or ^{35}S NMR data.

C14

67. (New) The method of claim 66, wherein the spectral data comprises calculated spectral data.

SUB D1

68. (New) The method of claim 57, wherein the calculated nuclear magnetic resonance data comprises ^{1}H, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{31}P or ^{35}S NMR data.

69. (New) A method for predicting a biological property of a test compound, comprising:

providing biological activity data for a plurality of compounds;

providing spectral data for the plurality of compounds;

providing spectral data for the test compound;

segmenting the spectral data for the test compound and the plurality of compounds into bins;

scaling the spectral data of the plurality of compounds in the bins;

weighting the spectral data of the plurality of compounds in the bins;

detecting a pattern of spectral data of the plurality of compounds in the bins that is correlated with the biological activity; and

detecting similarities between the pattern correlated with the biological activity and the spectral data of the test compound to determine if the molecule shares the biological activity.

70. (New) The method of claim 69, wherein scaling comprises autoscaling.

71. (New) The method of claim 70, wherein weighting comprises Fisher-weighting.

72. (New) The method of claim 69, wherein weighting comprises Fisher-weighting.

73. (New) The method of claim 69, wherein scaling comprises variance scaling.

74. (New) The method of claim 69, wherein the pattern detected in the bins comprises a set of canonical variate factors.

75. (New) The method of claim 69, wherein the spectral data comprises nuclear magnetic resonance, mass spectral, infrared, ultraviolet-visible, fluorescence, or phosphorescence data.

76. (New) The method of claim 75, wherein the spectral data comprises ^{13}C NMR data.

77. (New) The method of claim 76, wherein the ^{13}C NMR data comprises calculated ^{13}C NMR data.

78. (New) The method of claim 75, wherein the spectral data comprises a composite of two or more types of spectral data.

79. (New) The method of claim 78, wherein the composite comprises ^{13}C NMR data and EI-MS data.

80. (New) The method of claim 69, wherein detecting a pattern comprises statistical pattern recognition.

81. (New) The method of claim 76, wherein segmenting into bins comprises segmenting the ^{13}C NMR data into bins having a width from 0.5 ppm to 5.0 ppm.

82. (New) A computer implemented method for predicting the biological activity of a test compound, comprising:

receiving as input spectral data for a test compound;

receiving as input spectral data and biological activities of a training set of compounds;

segmenting the spectral data of the training set into bins;

autoscaling the spectral data of the training set;

Fisher-weighting the spectral data of the training set;

detecting a pattern of spectral data characteristic of an endpoint class of compounds in the training set of compounds;

segmenting the spectral data for the test compound into bins; and

predicting the biological activity of the test compound by detecting similarities between the segmented spectral data of the test compound and the pattern of spectral data characteristic of the endpoint class to predict whether the test compound is also in the endpoint class.

83. (New) The computer implemented method of claim 82, wherein detecting a pattern of spectral data associated with an endpoint class of compounds in the training set of compounds comprises calculating canonical variate factors for the bins.

84. (New) The computer implemented method of claim 82, wherein the spectral data of the test compound and the spectral data of the training set of compounds comprise spectral data selected from the group consisting of nuclear magnetic resonance data, mass spectral data, infrared data, ultraviolet-visible data, fluorescence data, phosphorescence data, and composites thereof.

85. (New) The computer implemented method of claim 84, wherein the spectral data of the test compound and the spectral data of the training set of compounds comprises a composite of two or more types of spectral data and the spectral data is normalized to yield structure descriptors of similar magnitude from each type of spectral data.